

# DSpace zase po roce (ale tentokrát s Vlastíkem)

Vlastimil Krejčíř, ÚVT MU

Otevřené repozitáře 2017

Brno, 7. 6. 2017

## Obsah přednášky

- DSpace 6
  - Přehled vlastností
  - Pár poznámek z praxe
- DSpace 7
  - Nahlédnutí pod pokličku

Přednáška bude techničtější, ale nelze jinak :-)

## DSpace 6

- Podzim 2016
- Aktuálně stále 6.0
  - Verze 6.1 by měla být do konce června
    - Opravy chyb (IP autentizace aj.)
- Zavádí řadu nových funkcí
  - Uživateli skrytých i viditelných
- ... a řadu starých funkcí ruší!

## DSpace 6: databáze (Hibernate)

- Zavedení frameworku **Hibernate**
  - Mapování objektů v programovacím jazyce do databáze
    - Usnadňuje práci vývojářům
  - Další vrstva API
  - Do budoucna snadněji možné napojit další databáze
    - Nyní podpora PostgreSQL a Oracle
- Od DSpace 5 i **FlywayDB**
  - Snadnější migrace databáze
- `$ bin/dspace database info`

## DSpace 6: UUID

- Vnitřně má každý objekt v DSpace UUID
- **Universally Unique Identifier**
  - Dlouhý identifikátor (128bitový)
  - Prakticky globálně jednoznačný
  - Např. *123e4567-e89b-12d3-a456-426655440000*
  - Dobré pro vývojáře :-) (i knihovníky)
  - Všimnete si toho v reportech

## DSpace 6: konfigurace (local.cfg)

- Lepší a přehlednější systém konfigurace
- Zaveden soubor **local.cfg**
  - Přepisuje nastavení z globální konfigurace
  - Stačí jen zapsat ty hodnoty, které chci měnit
  - => Tj. v praxi malý a přehledný
  - Změny v konfiguraci se projeví ihned!
  - A pro nahlédnutí zůstává i výchozí nastavení...

## DSpace 6: reporty (Healthcheck)

- Výpis a sumarizace informací o DSpace
  - Obecné - velikost assetstore a logs
  - Počet i velikost komunit, kolekcí, bitstreamů, ...
  - Kontrola checksums u bitstreamů
  - Agregace dat o uživatelích a skupinách
    - Mj. prázdné skupiny apod.
  - Analýza logů (počty *Exceptions*, *Warnings*, ...).

*Healthcheck by Josef Mišutka (UFAL UK). Díky!*

## DSpace 6: discovery (SOLR)

- Užití SOLR povinné
- Indexace souborů typu .x/s (MS Excel)
- Faseta umožňující filtrovat záznamy s alespoň jedním uloženým souborem
- Indexace textů psaných zprava doleva



## DSpace 6: vylepšení REST API

- REST API
  - Autorizace (LDAP, Shibboleth, ...)
  - Přístup do registrů (`/registries`)
  - Informace o DSpace (`/status`)

## DSpace 6: drobnosti

- Formuláře v JSPUI s možností výběru jazyka
  - Možno vložit např. popis anglicky i česky
  - Má to ještě mouchy...

Popis

český popis	Czech	Remove Entry
english description	English	Add More

## DSpace 6: drobnosti

- Podpora cloudového úložiště Amazon S3
- Podpora PubMed v XMLUI
- OAI-PMH v souladu s OpenAIRE 3.0
- Lepší bezpečnost (PG\_CRYPTO)

## To co bylo, dávno není...

- `build.properties`
- Starý browse systém (databázový)
- LNI modul (SOAP, RPC, ...)
- SRB (Storage Resource Broker)
- Nelze smazat skupiny *Administrators* a *Anonymous*
- Elastic Search Statistics modul
  - Doporučen SOLR nebo Google Analytics

## DSpace 6: něco z praxe

Upgrade (z DSpace 5.4) nešel úplně bez problémů:

- Přenos SOLR indexu
  - Instalátor říká vše OK
  - Pokus o přístup => Tomcat mrtvý (nutný `kill -9`)
    - Tj. nelze reindexovat
    - V logu nikde nic...
    - Pro neIT publikum, výše uvedené je **PEKLO**
  - **Řešení:**
    - Čistá instalace DSpace s přenesením databáze
    - Pečlivá kontrola konfigurace
    - `bin/dspace discovery-reindex`

## DSpace 6: něco z praxe II

Paměťové nároky nám trochu nabobtnaly...

- Už to běží, ale co to... nějak pomaaaalllllůůůůůů...
- Vytížení procesoru stoupá (Tomcat přes 200 %)...
- DSpace už nereaguje...
- ... a po nějaké době Tomcat (asi přes AJP) způsobí i pád Apache HTTP serveru :-(
- **Řešení:**
  - `-XX:MaxPermSize=256M -XX:PermSize=256M`
  - 128 MB je málo...

## DSpace 7: co nás čeká a nemine

- Podpora jen pro Java verze 8
- Opět vylepšení REST API
  - Plná podpora všech funkcí
  - Dle aktuálních *best practices*
- Jednotný systém řešení embarga

## DSpace 7: jednotné UI

- Nové uživatelské rozhraní
  - Místo XMLUI a JSPUI
- Odstranění schizofrenie a konec tříštění sil
  - Dělat vše 2x je vysilující
  - V oblasti PR to taky nevypadá dobře
- Zkoušeny různé technologie a přístupy
- Vítězem je **Angular 2**
  - Framework z dílny firmy Google



## DSpace 7: Angular

- Plně napojený na REST API
- Na bázi JavaScriptu
  - primárně používá nadstavbu TypeScript
- Angular Universal
  - Stránka je renderována na straně serveru
  - Prohlížeč dostává předpřipravené HTML
  - Vhodné pro indexátory (Google, Bing)

## DSpace 7: nové UI

- Co to bude znamenat?
  - Sám zatím neumím odhadnout...
  - Nicméně bude plné REST API, kdyby se to někomu nelíbilo...
- XMLUI a JSPUI již nebude od verze 7 vůbec podporováno



# Vaše dotazy?